

Федеральное государственное автономное
образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Институт филологии и языковой коммуникации
Кафедра журналистики

УТВЕРЖДАЮ
Заведующий кафедрой
_____ К. В. Анисимов
« _____ » _____ 2018г.

БАКАЛАВРСКАЯ РАБОТА

42.03.02 Журналистика

**ОПЫТ ПРИМЕНЕНИЯ ЖУРНАЛИСТИКИ ДАННЫХ В АВТОРСКОМ
МУЛЬТИМЕДИЙНОМ ПРОЕКТЕ**

Руководитель	_____	ст.преподаватель	Д.А. Устюжанина
Выпускник	_____		О.А. Ясева
Консультант	_____	канд.филол.наук, доцент	К.А. Зорин
Нормоконтролер	_____	ст.преподаватель	О.В. Богуславская

Красноярск 2018

СОДЕРЖАНИЕ

Введение.....	3
1 Журналистика данных как феномен.....	7
1.1 Понятия big и open data	7
1.2 Феномен журналистики данных.....	13
1.3 Методика работы с большими данными	18
1.3.2 Сбор данных.....	18
1.3.3 Обработка собранной информации.....	22
1.3.4 Визуализация.....	23
1.3.5 Публикация.....	27
2 Применение опыта журналистики данных в мультимедийном проекте «Задержите дыхание».....	28
2.1 Подготовительный этап работы над проектом.....	28
2.2 Сбор данных.....	29
2.3 Этап визуализации.....	33
2.4 Этап публикации.....	37
2.5 Анализ работы над проектом «Задержите дыхание».....	38
Заключение.....	43
Список использованных источников.....	45

ВВЕДЕНИЕ

В настоящее время под влиянием стремительного развития технологий возможности собирать, обрабатывать, хранить и распространять информацию становятся всё шире. Благодаря этому возникает тенденция к открытости данных: правительство и организации разного рода стремятся сделать свою деятельность прозрачной. Они публикуют в интернет-пространстве большое количество данных в машиночитаемом виде.

Для определения таких скоплений больших массивов данных в интернете появились термины «большие данные» и «открытые данные» («big data» и «open data» соответственно). Эти явления предоставляют простор для статистических и количественных измерений, которые позволяют увидеть определенную картину состояния общества, обнаружить те или иные проблемы в различных сферах жизни. В связи с тем, что зачастую массив данных столь огромен, что расшифровать его становится непосильной задачей для неподготовленного пользователя, в журналистике возникает новое направление – журналистика данных. Обработка и, главное, анализ таких массивов данных и их представление в удобном для читательского восприятия виде – важная и актуальная задача для современного журналиста.

Работая в этой области, журналист выступает в качестве дешифратора большого скопления информации, которую он должен преподнести аудитории в доступном и интересном виде – чаще всего в виде мультимедийного проекта, в который для наглядности включаются графики, карты, диаграммы.

Более того, журналистика данных находится на пересечении таких сфер как журналистика и программирование. С одной стороны, работа с данными, в первую очередь, включает в себя журналистскую составляющую: сбор информации, поиск источников, интерпретацию данных в читаемом виде. С другой – журналисту необходимо обладать навыками дизайнера и программиста, поскольку от него требуются умения извлечь цифровые данные

в различных форматах, уметь их обработать и представить в виде, интересном для читателя.

Зачастую исследователи называют журналистику данных рассказом историй с помощью цифр, а также поиском, анализом и визуализацией информации цифровых источников данных с целью создания уникального контента. Журналистика данных предполагает совмещение журналистом роли переводчика с «языка цифр» на язык, понятный всем и, одновременно, роли аналитика, представляющего глубокие выводы, сделанные на основе анализа данных.

Актуальность данной бакалаврской работы связана с постепенным распространением исследуемого явления в зарубежных и российских СМИ. Стоит отметить, что наибольшее развитие журналистика данных получила в Европе и США. Самыми успешными в этом направлении являются такие американские газеты, как The New York Times, Chicago Tribune, Los Angeles Times и другие. В России же журналистика данных еще не получила достаточно широкого распространения, однако мы уже можем встретить единичные примеры работы с большими данными как в самых крупных российских изданиях, так и в отдельных проектах активистов.

Проблема исследования заключается в том, что, несмотря на обилие открытых региональных источников данных, в региональной практике практически отсутствует журналистика данных как направление. Наша дипломная работа имеет практическую направленность и, исходя из проблемы, мы создали проект на экологическую тему, который, базируясь на открытых правительственных данных и открытых данных экологов, отвечает на несколько вопросов о красноярской экологии: где основные источники загрязнения воздуха, какие объемы выбросов происходят ежегодно, какова динамика ухудшения экологии и какие пути решения экологической проблемы в городе существуют.

Объектом исследования дипломной работы является специфика такого направления как журналистика данных в условиях работы с большими данными.

Предмет исследования – применение журналистики данных в региональной повестке на экологическую тему.

Цель работы – создать мультимедийный проект на экологическую тему, используя большие открытые данные. Исходя из цели, мы ставим следующие задачи:

- 1 Дать определение понятиям «открытые данные» и «большие данные», и выделить, в чем состоят их различия;
- 2 Выделить и описать этапы работы журналиста с открытыми данными;
- 3 Разработать концепцию собственного проекта;
- 4 Собрать и обработать открытые данные, найти другую необходимую информацию для проекта;
- 5 Создать собственный проект журналистики данных на экологическую тему;
- 6 Опубликовать проект в СМИ

В работе мы будем обращаться к таким теоретико-практическим источникам, как переведенное «РИА Новостями» «Пособие по журналистике данных», посвященное этапам работы с данными; прикладным советам по работе с большими и открытыми данными «The Guardian Datablog», статьям российского эксперта по открытым данным Ивана Бегтина. Также мы будем опираться на теоретические труды следующих исследователей: М. Г. Шилиной [Шилина, 2009, 2011, 2013], С. И. Симаковой [Симакова, 2013, 2014], [Майер-Шенбергер, Кукьер, 2014], П. А. Клеменкова, С. Д. Кузнецова [Клеменков, Кузнецов, 2012], М. Н. Шерстюковой [Шерстюкова, 2012], В. Ю. Левченко [Шилина, Левченко, 2014], А. Моррисона [Моррисон, 2010], С. А. Панюковой [Панюкова, 2015], С. Горансона [Goranson, 2014].

Методы исследования: индукция, дедукция, анализ, сравнение.

Исходя из поставленных задач, мы предлагаем следующую структуру исследования: первая глава бакалаврской работы посвящена определению понятий «большие данные» (big data), «открытые данные», «журналистика данных», «мультимедийный проект». Также в первой главе мы даём авторское определение журналистики данных и описываем этапы работы над проектом журналистики данных. Во второй главе исследователь описывает собственный опыт создания экологического проекта в области журналистики данных, подвергает анализу трудности, с которыми столкнулся при работе с данными, сопоставляя теоретический план работы с данными и то, что получилось на практике. В заключении подведены общие итоги исследования.

1 Журналистика данных как феномен

По подсчётам International Data Corporation, мировой рынок больших данных неуклонно растёт: в 2015 году он достиг \$122 млрд, в 2016 году – уже \$130 млрд. К 2020 году аналитики прогнозируют рост объема рынка до \$203 млрд [Отчет ICD, 2016]. Кроме того, все больше государственных и негосударственных структур, стремясь к прозрачности деятельности, выкладывают данные в открытый доступ. Однако зачастую эти данные настолько не структурированы и хаотичны, что из них невозможно ничего понять. Именно в этот момент начинает свою работу журналистика данных. Журналист в этом случае выступает посредником между организацией и обществом, используя открытые большие данные как ключ к пониманию общественных/ экономических/ политических вопросов.

В первой главе мы дадим определения понятий «открытые данные» и «большие данные», работа с которыми составляет основу журналистики данных. Далее мы определим специфику феномена журналистики данных и выделим этапы работы журналиста с данными, подробно описав каждый из них.

1.1 Понятия big и open data

Термин «большие данные» (big data) возник десятилетие назад. В 2008 году его ввел в употребление редактор научного журнала Nature Клиффорд Линч, рассказывая о росте объемов информации во всем мире, который происходит ежемоментно [Big data: How do your data grow, 2008]. Эти объемы информации скапливаются в огромные массивы – базы данных, которые не просто хранят информацию о чем-либо, но и позволяют извлекать ее.

В настоящий момент данных, которые обретают цифровую форму, всё больше. Их объемы растут, накапливаются, этот процесс непрерывен и

необратим. Более того, международные, государственные и негосударственные организации и правительство, стремясь к открытости и прозрачности, делают эти массивы данных открытыми. Такое явление и называется «открытые данные».

В журналистской практике работать с большими данными приходится все чаще, и это требует от журналиста особых навыков. Так, ему требуется совмещать в себе сразу несколько ролей: журналиста, социолога, программиста и дизайнера. Для журналистов базы данных становятся основой исследования социальных, политических и экономических проблем и явлений. Более того, зачастую журналист – главный посредник между разного рода организациями и физическими лицами, способный интерпретировать большие массивы данных и донести их до аудитории, потому для нас актуально определить специфику больших и открытых данных. В этом параграфе мы рассмотрим явления больших и открытых данных, определив их специфику и роль в работе журналиста в направлении журналистики данных.

Чтобы лучше понимать явление больших данных, для начала рассмотрим несколько его определений. В книге «Большие данные: революция, которая изменит то, как мы живем, работаем и мыслим» Виктор Майер-Шёнбергер и Кеннет Кукьер подчёркивают, что для термина «большие данные» не существует единого определения [Шёнбергер, Кукьер, 2014, с. 45]. Авторы пишут, что благодаря росту информации, не помещающейся в обычный компьютер, стали появляться новые инструменты для её извлечения, именно это и характеризует в первую очередь большие данные. Также авторы указывают на то, что это операции, которые можно выполнять только в большом масштабе.

Компания International Data Corporation (IDC), исследующая мировой рынок информационных технологий и телекоммуникаций с 1964 года, в своем исследовании определяет этот термин так: «Большие данные – новое поколение технологий, предназначенных для экономически эффективного извлечения полезной информации из очень больших объемов разнообразных данных путем

высокой скорости их сбора, обработки и анализа» [Большие данные в финансовой отрасли: обзор и оценка перспектив развития мирового и российского рынков, 2016, с. 5]. В данном определении большие данные рассматриваются в наиболее современном их понимании: не только как скопления информации, но, в первую очередь, как инструменты, с помощью которых можно извлекать и анализировать эти данные, получая выгоду. В исследовании мы будем опираться, прежде всего, именно на такое понимание big data. Также за рубежом «большими данными» называют «совокупности данных с возможным экспоненциальным ростом, которые слишком велики, слишком неформатированы или слишком неструктурированы для анализа традиционными методами» [Поиск эффективных инструментов работы с большими данными, 2010, с. 41-50].

В российских исследованиях термин рассматривается уже. М. Шилина определяет термин «большие данные» как наборы данных, «размер которых не выходит за пределы возможностей типичных программных средств сбора, хранения, управления и анализа баз данных» [BigData, OpenData, LinkedData, метаданные в PR: актуальные модели трансформации теории и практики, 2013, с. 3]. С этим термином стоит поспорить, так как некоторые исследователи указывают, что для извлечения и исследования крупных массивных данных помимо типичных инструментов (таких, как Excel, например) требуются и такие, для использования которых у обычных пользователей не хватает навыков.

Мы совместим определения зарубежных ученых, определив большие данные как крупные скопления массивов данных, для извлечения которых требуются особые инструменты и методики. Таким образом, мы подчеркнём не только наличие самих данных, но и необходимость специфических инструментов для работы с ними.

Большие данные характеризуются через «три V»: Volume, Variety, Velocity (объем, разнообразие и скорость соответственно) [Основы Data Science, 2016].

Исходя из понятия, можно предположить, что объем – самый важный признак больших данных. Важно также понимать, что эти «большие» массивы данных неструктурированы, хаотичны и представлены в различных форматах: они могут быть открытыми или закрытыми, могут требовать разной технологии обработки, хранения и передачи. Тот факт, что количество данных, производимых как людьми, так и машинами, стремительно растет, заставляет ИТ-инфраструктуру постоянно совершенствовать систему хранения, обработки и предоставления доступа к данным [Основы Data Science, 2016].

Под скоростью в «big data» понимается не только скорость, с которой данные поступают, но и скорость, с которой эту информацию можно извлечь. Кроме этого, предполагается, что информация быстро обновляется (например, «поточные данные» обновляются в режиме реального времени - запросы в поисковой системе, показания датчиков и т.п.).

Третий признак – разнообразие – указывает на то, что данные содержат различную по типу информацию (структурированную и неструктурированную), представленную разными организациями. Данные могут быть представлены в виде текста, аудио-, видео- форматов, данных карт памяти, данных со спутниковых и навигационных систем, информации из социальных сетей и поисковых запросов [Феномен Big Data, 2014].

Нередко эти три параметра дополняют четвёртым – Veracity (достоверность), характеризующим представленных данных.

Источниками больших данных могут выступать:

- социальные сети и статистика из них;
- сводки с измерительных устройств (спутники, датчики);
- журналы доступа пользователей веб-сайтов;
- научные сводки;
- базы данных различных ведомств и организаций

Область применения больших данных широка: они универсальны и могут быть использованы в различных сферах деятельности, начиная от энергетики, горнодобывающей и нефтяной промышленности и заканчивая банковскими

системами и сферой здравоохранения. Ценность больших данных состоит в том, что они позволяют обозначить или предсказать проблемы, которые возникают или могут возникнуть в обществе или помогают оптимизировать различные общественные, государственные, бизнес-процессы.

Явление больших данных тесно связано с открытыми данными. Термин «открытые данные» (open data) приобретает все большую популярность и связано его распространение в большей степени с правительственным желанием открытости своей деятельности и соответственно публикации открытых данных о ней.

Чтобы более подробно изучить явление открытых данных, дадим определение понятию. В «Open data handbook» этот термин раскрывается так: «Открытые данные – это информация, которую кто угодно может свободно использовать и распространять. Допустимы лишь требования указывать источник данных и распространять их на тех же условиях, что и исходные <...> Каждый имеет возможность использовать и распространять данные» [Open data handbook, 2012].

На портале административной реформы содержится следующее определение открытых данных: «Это информация о деятельности государственных органов и органов местного самоуправления, размещенная в сети Интернет в виде массивов данных в формате, обеспечивающем их автоматическую обработку для повторного использования без предварительного изменения человеком (машиночитаемый формат), и на условиях ее свободного (бесплатного) использования» [Портал административной реформы].

В работе будем использовать второе понятие, поскольку оно содержит акцент на публикации именно государственных данных, чего нет в определении из «Open data handbook» (однако будем иметь в виду и то, что не только государство публикует открытые данные, но и частные компании – этого ни в одном определении нет).

Среди базовых характеристик открытых данных выделяют объем, скорость прироста, обработки, получения результатов и многообразие структурированных и полуструктурированных данных. Важно, чтобы открытые данные были доступны и читаемы, позволяли всем пользователям повторно их использовать и бесплатно распространять.

Значимость открытых данных во многих отраслях общественного и государственного строев колоссальна. Благодаря открытым данным обеспечивается прозрачность работы государственного аппарата и его работы с налогами населения, формируется база для гражданского контроля, появляются новые услуги и условия работы для бизнеса и граждан. Однако в связи со слабой информированностью, с цифровым разрывом и недостатком технологических компетенций как у граждан, так и у властей, на мировом и российском рынке в частности складывается ситуация слабого использования открытых данных.

В отличие от открытых данных, big data имеют большой коммерческий потенциал. М. Шилина говорит о том, что «большие» персональные данные, данные о покупках и денежных тратах потребителя могут быть монетизированы [Шилина, 2015].

По словам М. Шилиной, в работе с большими и открытыми данными большую важность приобретает способ их обработки и представления, то есть «аналитическая компонента, в первую очередь, профессиональная» [Шилина, 2015, с. 12]. Это означает, что для журналиста становятся важными навыки аналитика, что важнее – программиста и дизайнера, он должен научиться совмещать их или, как альтернатива, начать работать в команде с такими профессионалами.

Говоря о современных тенденциях в области больших данных, Кукьер и Шененберг подчеркивает, что в работе с массивами информации беспорядочность и гибкость преобладают над точностью, и, как ни парадоксально, этот тренд придает итоговому проекту качество и результативность [Революция, которая изменит то, как мы живем, работаем и

мыслим, 2014]. Выходит так, что с увеличением масштаба информации и скорости развития технологий точность в представлении данных не имеет такого значения, как прежде.

Мы рассмотрели понятие больших и открытых данных и обозначили их свойства и актуальность. Таким образом, пришли к выводу, что работа с big data помогает исследовать различные сферы жизни: от политика до частных человеческих проблем. Следующим шагом важно исследовать инструмент работы с массивами информации – журналистику данных. В следующем параграфе мы рассмотрим это явление, благодаря которому огромные базы данных становятся историей, рассказывающей читателю о том или ином явлении.

1.2 Феномен журналистики данных

Благодаря возникновению новых форматов данных, стали появляться особые методы работы с ними. Рождается такое явление, как журналистика данных (от англ. «data journalism»). Протоформой журналистики данных является так называемая прецезионная журналистика, возникшая в США в 70-е годы [Пособие по журналистике данных, 2012]. Прецезионная журналистика отличалась предельной объективностью и применяла научно-исследовательские методы: опросы, статистические исследования и другое.

Журналистика данных перенимает аналитическую компоненту прецезионной журналистики и раскрывает ее в новых форматах, которые предоставляют современные технологии. Также прямой предшественницей журналистики данных была «компьютерная журналистика», которая стала первым организованным, систематическим подходом к использованию компьютеров для сбора и анализа данных для улучшения новостей [Пособие по журналистике данных, 2012].

Об отличии журналистики данных от других видов журналистики исследователь в сфере дата-журналистики Пол Брэдшоу пишет: «Возможно, это новые возможности, которые открываются, когда вы объединяете традиционный «нюх на новости», умение выведать все, что случилось, со способностью рассказать захватывающую и наглядную историю, с настоящим масштабом и разнообразием цифровой информации, которая ныне доступна» [Пособие по журналистике данных, 2012, с. 3].

Поскольку в большей степени журналистика данных – это совокупность нескольких профессий, при попытке определить границы понятия возникают трудности. М. Шилина определяет термин таким образом: «Журналистика данных – это набор специфических навыков для поиска, анализа визуальной информации цифровых источников данных с целью формирования интерактивных форматов уникальной подачи авторского аналитического журналистского контента и эффективного взаимодействия СМИ, журналиста с аудиторией; это формат актуального журнализма; формат профессионального журналистского медиаконтента: метод его создания, трансляции, потребления» [Шилина, 2014]. В данном определении, по нашему мнению, автор несколько опускает первичность функции общественного контроля, а цель авторского самовыражения для журналиста ставит выше, а потому определение М. Шилиной не в полной мере отражает сущность журналистики данных.

Более неформально определяет данный термин И. Бегтин: «Журналистика данных – это совмещение журналистом роли переводчика с экономического/государственного цифрового языка в язык понятный всем и, одновременно, роли аналитика, представляющего материалы в предельно наглядной форме» [Информационная культура, 2014]. По словам И. Бегтина, это могут быть списки, графики, интерактивные карты, которые позволяют отображать данные разными способами. Во всех случаях текст является вспомогательным, а не основным инструментом. Данное определение возьмем за рабочее, поскольку в нем отражена как специфическая роль журналиста в журналистике данных, так и инструменты, при помощи которых можно

работать с данными. Однако в определении понятия не хватает упоминания факта работы журналиста именно с большими открытыми данными, и при использовании понятия «журналистика данных», будем учитывать этот аспект.

Первая и ключевая причина того, почему популярность журналистики данных набирает обороты – это наглядность способа представления больших объемов информации: графики, таблицы и пр. Подобные материалы привлекают внимание аудитории, которая не всегда готова вчитываться в большие тексты, но хорошо воспринимает их визуальное представление. С одной стороны, журналистика данных стремится облечь большие массивы данных в наиболее наглядную и удобную для восприятия форму, но с другой – усложняются процессы подготовки материалов – необходимо извлечь данные, обработать их и визуализировать. Это говорит о тенденции перехода к тому типу производства и потребления медиа, когда анализ преобладает над скоростью, а также когда аудитории предоставляется возможность дополнять созданный журналистом проект на основе открытых данных.

Говоря о функциях журналистики данных, Арон Пилхофер, журналист New York Times, говорит о возможностях создавать новые сюжеты: «Она может включать в себя все – от традиционных репортажей, которые готовятся при помощи компьютера (с использованием данных в качестве «источника») до самых современных и передовых способов визуализации данных и новостных приложений.<...> Объединяющая цель – журналистская: предоставление информации и анализа, чтобы помогать информировать нас обо всех важных проблемах дня» [Пособие по журналистике данных, 2012].

Функция социального исследования заключается в том, что журналистика данных в некотором смысле становится «первичным исследователем статистических данных», занимается мониторингом отчетов, опросов. В данном случае дата-журналистика делает общественно важные прогнозы в жизни социума и выявляет закономерности и проблемы социума.

Основываясь на результатах анализа открытых данных, журналистика данных реализует функцию общественного контроля СМИ за социальными

институтами. Таким образом, она указывает на проблемы в разных сферах общества, привлекая внимание общества к ним и предлагая пути ее решения. Р. Жолудь отмечает, что эта функция особо задействована при применении журналистики данных в журналистских расследованиях. К примеру, расследование The Telegraph Group в 2009 году, использующее открытые данные и разоблачающее слишком высокие расходы членов парламента Великобритании, вызвало большой политический скандал [Журналистика данных: предпосылки возникновения, функции, возможности, 2015].

Следует отметить, в журналистике данных меняется процесс работы над материалами: используются новые технологии, которые не применяются в традиционных типах СМИ, такие как программное обеспечение для сбора, очистки, анализа и визуализации данных. Кроме того, меняются этапы работы журналиста с данными. М. Шилина в статье «Data Journalism – дата-журналистика, журналистика метаданных...» [Шилина, 2014]. Этапы работы с данными делит на те, что относятся к субъекту – то есть к журналисту – , и те, что относятся к объекту – аудитории. К задачам журналиста М. Шилина относит нахождение темы, информации, далее – анализ с помощью специальных приложений и программ, также журналист должен визуализировать добытую информацию и интересно для аудитории представить статистические данные в форме сюжета или проекта [Шилина, 2014]. К функциям аудитории – дополнение представленных в открытом доступе проанализированных данных.

Если говорить о способах визуализации, то стоит упомянуть, что журналистский материал в дата-журналистике – это, по большей части, совокупность онлайн-журналистского текста, основанного на больших данных, и представления этих данных в разных формах: таблицы, карты, диаграммы, инфографика и т.п. Обновление проекта зачастую происходит онлайн: либо по мере поступления новых данных, либо в процессе подключения читателей к модераторству текста.

М. Шилина говорит, что тематика материалов, основанных на открытых данных, как и в традиционной журналистике, не имеет ограничений: начиная

бизнесом, доходами политиков, заканчивая социально-экономическими аспектами вроде ЖКХ, уровня жизни населения и т.п. [Шилина, 2014]. Важно то, как тема раскрывается с помощью больших данных: она приобретает глубину и объективность.

Если говорить об ограничениях в журналистике данных, то многие исследователи главной трудностью для журналиста называют необходимость уметь работать с рядом таких аналитических и технических инструментов, как Excel, Google Docs, IBM ManyEyes, Wordle и многие другие. В данном случае журналист выступает с позиции программиста (или наоборот) и в обеих ситуациях необходимо учиться методам работы с информацией. Именно по той причине, что интерпретировать полученные данные следует особенно тщательно, обычно в штате крупнейших международных деловых газет, вроде Financial Times, нет ни одного человека, чья должность звучала бы как «журналист по работе с данными». Над проектами, связанными с обработкой больших массивов информации, работают коллективы специалистов из разных областей. Обычно в рабочую группу входят специальный корреспондент, знакомый с изучаемой проблемой, пишущий журналист, специалист по интерактивной графике, дизайнер и зачастую программист. Именно потому, что журналистика данных является совокупностью профессий, по словам И. Бегтина, в России не так много журналистов, готовых создавать проекты в рамках данного направления.

1.3 Методика работы с большими данными

А вторы книги «Основы data science» называют следующие этапы работы с данными: постановка цели проекта, сбор данных, их подготовка и исследование; после чего моделирование, или визуализация, и в случае журналиста это продвижение проекта. Практики, работающие с большими данными, указывают на то, что такой структурированный подход повысит

эффективность в работе, однако это немного утопическая модель: в реальности нужно будет снова и снова возвращаться на тот или иной этап и исправлять некорректные данные. Несмотря на то, что в реальности такому плану следовать трудно, нужно помнить, что заранее заданный алгоритм действий поможет прийти к четкому результату еще до получения самих данных.

Кроме того, в пособии подчеркивается, что составление плана поможет делегировать работу. «Невозможно быть специалистом во всех областях сразу», – напоминает он.

1.3.1 Подготовительный этап

Закономерно, что первым шагом в работе над любым проектом – не только журналистики данных – нужно составить план. Но, что важно, в случае с большими данными план поможет структурировать и направить рабочий процесс, сложный и несколько хаотичный.

Прежде чем начать сбор данных, журналист должен продумать вопросы, на которые в итоге хочет получить ответы.

Журналист в данном случае работает как бы в обратном порядке: сначала он формулирует итоговое утверждение, а после – подкрепляет их данными. После постановки цели и вопросов нужно определить, какие переменные и данные необходимо найти и переработать, чтобы прийти к тем утверждениям.

1.3.2 Сбор данных

На этом этапе журналист должен быть готов к тому, что данные могут быть как структурированными, так и напротив – хаотичными. В предыдущем параграфе мы подчеркнули, что вероятность встретиться с разрозненными данными куда выше, и бояться ее не стоит. Исследователи делят данные на структурированные и неструктурированные. Чаще всего мы можем встретить

данные в виде таблиц Excel и CSV: информация в них представлена в форматах XLS, XLSX. Также часто используется формат CSV, где данные отделяются друг от друга запятой или точкой с запятой. Чаще всего, если журналисту попались данные в виде таблиц, будет означать, что работа с ними не потребует сильного форматирования или очистки.

Еще один формат данных – PDF, который зачастую включает в себе структурированные таблицы, собранные на компьютере, и не позволяющие копировать данные из этих таблиц. Для работы с таким форматом журналисту необходимо воспользоваться специальным программным механизмом, извлекающим данные из этого формата.

Больше всего затрудняет работу журналиста с данными формат отсканированных изображений – например, PNG. Это чаще всего изображения, которые считываются компьютером как цельный огромный блок. Для обработки таких файлов понадобятся программы оптического распознавания текста.

Еще один вид представления данных – неструктурированные. Это могут быть данные, оформленные в виде текста и требующие вмешательства программиста: ему понадобятся инструменты программирования Python или Ruby для «скрейпинга» данных с помощью `morph.io`. Немного подробнее остановимся на «скрейпинге», или, как этот процесс называют авторы «Пособия по журналистике данных», – «выскребании». Этот процесс помогает извлечь структурированные данные из веб-страницы с помощью утилиты или небольшой части кода [Пособие по журналистике данных, 2014, с. 19]. «Скрейпинг» необходим в случае, если перед журналистом находятся машиночитаемые данные – то есть информация, предназначенная для компьютерной обработки, а не для пользования человеком. Машиночитаемыми данными является, например, форматы CSV, XML, JSON и XSL, XSLX, при этом такие форматы, как, например, doc, страницы HTML и документы PDF предназначены скорее для визуального представления информации. «Скрейпинг» нужен в случаях, когда не удастся скопировать таблицу из

интернета в формат Excel или данные хаотично расположены на разных страницах в интернете и собирать информацию с каждой из них самостоятельно долго и трудно. Именно для автоматизации процесса сбора и существует «выскребание» – использование специального кода, который можно применить практически на любом сайте. Для того, чтобы «выскрести» данные, можно использовать следующие инструменты: Readability, ScraperWiki (извлекают из веб-страницы текст), DownThemAll (загружает несколько файлов одновременно). Также можно скачать специальные приложения для браузеров: для Chrome – Scraper extension, FireBug – для Firefox и другие. Крупнейшими сервисами, «которые помогут встроить коды «скребков» в программы, написанные на различных языках, в т. ч. на Python, Ruby и PHP. Произвести сбор данных с других веб-сайтов вам могут также помочь такие сервисы, как Google Spreadsheets and Yahoo! Pipes».

Непопулярным форматом представления данных можно считать данные, сохраненные пакетами, которые можно отправить на веб-сайты или обработать с помощью статистического программного обеспечения, например, JavaScript Object Notation или формат XML. Чтобы воспользоваться такими данными, их нужно преобразовать в CSV или Excel формат с помощью, например, сайта konklone.io.

Источники данных совершенно не ограничены: ими могут стать как таблицы с сайта правительства, так и живые люди. Ниже мы рассмотрим, где журналист может найти нужные ему большие данные.

Один из наиболее простых способов обратиться к данным – расширенный поиск Гугла. В него включены несколько полей для поиска и таблица, в которой есть функции расширенного поиска.

Также одним из главных источников для журналиста могут стать официальные данные, взятые, например, с государственных сайтов или сайтов правительств страны или регионов. В российской практике нередко можно столкнуться с тем, что раздел «открытые данные» есть, однако полезной

информации журналист там едва ли найдет, поскольку правительство выкладывает информацию зачастую «для галочки».

Одной из ключевых площадок больших открытых данных не только в мире, но и в России можно считать data.gov.ru/opendata, а также сайты правительств и администраций. Также, если говорить о зарубежных источниках, удобным сайтом выступает Guardian World Government Data - поисковик, в котором можно найти каталоги баз данных, публикуемых иностранными правительствами. Также автоматизированным ресурсом поиска открытых данных является The Data Hub, поддерживаемый фондом «Открытая информация» (Open Knowledge Foundation), который способствует легкому поиску, обмену и использованию открытых доступных источников данных. Для поиска информации за несколько лет и обо всех странах можно обратиться на порталы данных Всемирного банка, Организации Объединенных Наций, Всемирной Организации Здравоохранения, данных UNICEF и т.д.

Обмен данными и перепродажа осуществляются на порталах Buzzdata, Infochimps, и DataMarket. Вокруг этих сайтов построились целые сообщества тех, кто готов обмениваться данными не только бесплатно, но и за деньги.

Своеобразной социальной сетью больших данных выступает сервис Get The Data. На этом сайте можно задать вопросы и ответы по поводу баз данных: где найти ту или иную информацию, какими инструментами его извлечь и т.п.

Мало популярным в России и неочевидным для нас способом получения данных являются рассылки. Например, Data Driven Journalism List и the NICAR-L, рассылки фонда Open Knowledge Foundation могут стать полезным источником информации – ведь данные вы получите от таких же журналистов, погруженных в тему биг-дата и занимающихся проектами. Есть вероятность того, что кто-то из них делал схожий с вашим проект и сможет предоставить нужную информацию.

Еще одним сообществом людей, работающих с данными, выступает Hacks/Hackers. У этого ресурса есть преимущество: на этой площадке

объединились журналисты и программисты, что вдвое увеличивает шанс на то, чтобы найти нужные данные.

Эксперты-практики советуют также обращаться и к живым источникам. Например, журналист может спросить госслужащего или эксперта в той или иной области, где можно найти ту или иную информацию.

Когда вышеперечисленные источники данных недоступны, другие можно собирать методом краудсорсинга, когда аудитория сама предоставляет информацию о себе или рассказывает о том, свидетелем чего стала. Так, по крупицам, небольшая информация об одном человеке собирается в большие данные о целой стране. Еще один вид «живого» источника – эксперт, который поможет разъяснить собранные данные.

1.3.3 Обработка собранной информации

Для того, чтобы правильно интерпретировать данные и прийти к выводам, информация должна пройти процесс стандартизации – то есть, очистки данных. Это нужно для того, чтобы исключить ошибки, повторяющиеся числа и т.п.

Первое, что необходимо сделать на этом этапе – убедиться, что столбцы в таблице обозначены верно и что тип данных в строках соответствует заголовку. Например, если таблица состоит из таких колонок, как имя, возраст, единицы измерения или дату, нужно привести эти показатели в единый вид, чтобы при компьютерной обработке не возникло ошибки. Например, при работе с датой, можно подвести все данные под формат ДД/ММ/ГГГГ, тогда информация будет считана верно.

Также при стандартизации журналист должен проверить, чтобы буквы были одного регистра – заглавные или строчные, а в строках нет лишнего лишних пробелов.

При затруднениях поможет программа Open Refine, способная привести к единому формату даже самые запутанные наборы данных.

1.3.4 Визуализация

Авторы «Пособия по журналистике данных» пишут, что визуализация – «рабочая лошадка» журналистики данных. Во-первых, на этом этапе можно, благодаря визуализациям, понять, какие темы можно исследовать далее в репортаже. Кроме того, эта ступень показывает отклонения в данных – присутствуют ли какие-либо резкие отклонения, пробелы, ошибки или, напротив, наиболее удачные «истории».

Для аудитории визуализации имеет непосредственное значение: читателю нагляднее представлены огромные данные, он может их сопоставлять и исследовать, возвращаться к ним.

Журналисты-практики призывают не считать это отдельным шагом, который нужно выполнять после того, как весь материал собран. Нужно начинать оформлять работу визуально, как только вы добыли часть информации – тогда проще будет, как мы сказали ранее, отследить недочеты или нехватку данных.

Как мы говорили ранее, большие данные – способ рассказать историю. Поскольку все массивы данных обрабатываются на компьютере, и для удобства читателя журналист делает визуализацию интерактивной, логично предположить, что итогом работы с большими данными должен стать онлайн-проект, если точнее – мультимедийный лонгрид. Мультимедийный лонгрид может вмещать в себя текстовые вставки, «скачки» (нумерованные списки, отчеты, ранжирование, рейтинги, очки и др.), «цепи» (образцы, инструкции, временные шкалы), карты (географические, геологические, статистические), схемы, диаграммы [Голомбински, Хаген, 2013].

Конвергенция жанров и форм на онлайн-платформах привела к появлению мультимедийного текста, совмещающего в себе мультимедийный проект и проект журналистики данных. По своему определению, мультимедиапроект включает в себя текстовое, фото-, видеосопровождение, а также инфографику. Также мультимедийный текст может обладать интерактивностью, позволяющей пользователю «взаимодействовать» с контентом. М.Н. Булаева в статье «Мультимедийный лонгрид как новый журналистский формат» добавляет, что лонгрид подразумевает также «повествование, посвященное актуальной социально значимой теме, имеющее драматическую структуру, сочетающее в себе свойства журналистики и художественной литературы, созданное на основе сочетания различных медиаплатформ (текст, фото, видео, аудио, графика, инфографика, анимация) [Мультимедийный лонгрид как новый журналистский формат, 2015]. Таким образом, проект журналистики данных, представленный в виде лонгрида, включает в себя все форматы и качества привычного онлайн-журналистского проекта.

Как и в работе со стандартным мультимедийным текстом, в проекте журналистики данных важную роль играют инструменты визуализации, поскольку каждый из них предназначен для определенной цели. Горансон и его коллеги в своей работе «Датаблог...» выделяют следующие основные виды визуализаций: таблицы, карты, графики, диаграммы, схемы [Goranson C. Data Visualization For Big Data, 2014].

Один из наиболее удобных форматов работы с большими данными – таблицы. Они служат для того, чтобы показать определенные данные в виде столбцов и колонок. Плюсы таблицы заключаются в том, что, во-первых, пользователь может сам интерпретировать данные, и, во-вторых, большинство пользователей хорошо знакомы с таблицами и умеют их читать. Недостатки так же можно отнести к пользовательским ожиданиям: если пользователь не сможет изучать таблицу как ему удобно (например, она начнётся с неважного для него параметра, или он не сможет увеличить ширину столбца), то он не

будет с ней разбираться. Авторы «Пособия...» рассматривают этот способ оформления данных как наиболее наглядный способ представить небольшие данные [Пособие по журналистике данных, 2014]. Таблицы придают данным организованный и фильтрованный вид. Главным ограничением таблиц журналисты-практики видят бесполезность в работе с несколькими показателями одновременно (например, изменение численности населения страны в течение времени).

Следующий вид визуализации – графики (graphs). Авторы «Датаблога...» называют их «традиционно тяжёлыми для восприятия», поскольку они требуют от пользователя достаточно много времени и внимания. Как правило, графики несут в себе очень много информации, зачастую – перегружены ею. Основным преимуществом графиков авторы называют универсальность, недостатком – как сказано выше, перегруженность информации, которая затрудняет восприятие информации. Среди основных видов графиков перечислим следующие: линейный, график рассеивания.

Диаграммы (charts) – еще один способ визуализировать информацию. Диаграммы визуализируют количественные и качественные связи. Разные виды диаграмм используют для демонстрации данных, качественных и количественных связей, включают в работу над информацией пространственное мышление в дополнение к логическому. Их преимущество заключается в лёгком представлении сравнений и пропорций. Но переполнение диаграмм информацией делает их зачастую непонятными. Среди основных видов диаграмм выступают столбчатые диаграммы, гистограммы, тепловая, кольцевая, лепестковая диаграммы.

Схемы – также достаточно популярный способ наглядно представить большие данные, поскольку они дают возможность показать «всю вселенную данных» в одном графике. Диаграммы изображают связь двух или более субъектов и последовательно показывают поток информации через сеть связанных между собой узлов. На диаграммах удобно обозначать географические связи между данными, например, изменение статистики в

направлении с севера на юг или от городов к сельской местности. Подобный способ визуализации зачастую может быть чрезмерно нагроможденным, что вызовет затруднение при его чтении.

Способы визуализации, по словам авторов «Пособия по журналистике данных», играют различные роли:

1 Они могут проиллюстрировать какой-то момент сюжета более наглядным способом;

2 Могут убрать из повествования ненужную техническую информацию;

3 Особенно в тех случаях, когда они интерактивные, они могут давать возможность осуществлять исследовательские работы, а также обеспечивать прозрачность в отношении вашей репортажной работы в глазах читателей

На этом этапе журналисты The Guardian Datablog советуют не прибегать сразу к инструментам визуализации, а вначале провести «мозговой штурм» с командой дизайнеров. Специалисты The Guardian обозначают проблему, которая заключается в том, что многие сервисы предполагают ограниченное количество способов визуализации, отчего графики могут быть неточными, недостаточно информативными или иметь лишние детали, которые изначально проект не предполагал [The Guardian Datablog, 2015].

Основные инструменты журналистов данных:

- электронные таблицы: LibreOffice, Excel или Google Docs
- геоинформационные системы (GIS): Quantum GIS, ArcGIS, GRASS
- библиотеки для визуализации: d3.js (mbostock.github.com/d3), Prefuse (prefuse.org), Flare (flare.prefuse.org)
- инструменты обработки данных: Google Refine, Datawrangler
- программы для визуализации без программирования: ManyEyes, Tableau Public (tableausoftware.com/products/public)

В дальнейшем для создания графиков и диаграмм мы планируем использовать такие онлайн сервисы, как Infogram, Piktochart, Create.ly и т.д.

Третий этап — интерактивная визуализация данных. Это способ графического представления информации, который позволяет читателю взаимодействовать с системой отображения информации и наблюдать ответную реакцию системы. Данный способ используется для анализа изменений тенденций, взаимосвязей, соотношений и закономерностей в исследуемом наборе однородных параметров предметов или явлений. Данную технологию еще называют интерактивной инфографикой [Шиманский, 2012]. Инструменты интерактивности позволяют пользователям анализировать данные, взаимодействуя с «визуальной репрезентацией данных», а также с цветом и формой визуализации [Gartner IT Glossary].

Инфографика может быть представлена тремя формами: статическим или движущимся изображением и интерактивным интерфейсом. Проект журналистики данных всегда представлен интерактивной инфографикой с интерактивным интерфейсом.

1.3.5 Публикация

Этот этап состоит не только из самой публикации. Авторы «Пособия...» сообщают, что получить отзыв читателя — не менее важная часть постработы с проектом. Аудитория может проверить или добавить факты, приведенные журналистом.

Но важно, что аудитория — это не только размытый образ, случайный читатель, но и вполне конкретный целевой пользователь — нужно исходить из заданной журналистом темы и искать связанных с ней людей. «Существуют ли какие-либо торговые организации, с которыми вы можете вступить в контакт, и которые могут изъявить желание обратить внимание общества на вашу работу и прорекламировать те источники, которые вы получили?», — такие вопросы должен задавать себе журналист в ходе продвижения проекта [Пособие по журналистике данных, 2014].

Одним из способов обратной связи и дополнения уже готовой работы журналисты авторы пособия считают присоединение комментариев к полям в таблице или к инфографике.

Важно, чтобы проект имел отклик и благодаря обратной связи журналист мог оценить, насколько полно и понятно он донёс данные, было ли это интересно и понятно аудитории.

2 Применение опыта журналистики данных в мультимедийном проекте «Задержите дыхание»

«Задержите дыхание» – мультимедийный проект об экологической обстановке в Красноярске. Основой для проекта послужили базы данных министерства экологии Красноярского края, сайта правительства Красноярского края, сайта администрации города Красноярска, Среднесибирского УГМС, Красноярскстата, проекта «Красноярск. Небо».

Точка доступа: <http://zaderzhite.dyhanie.tilda.ws/>

Цель проекта – ответить на главные волнующие красноярцев вопросы об ухудшении экологической обстановки в городе: почему так происходит, что делает правительство края для изменения ситуации и когда горожанам ждать улучшений.

Задачами проекта стали сбор воедино многочисленных данных о состоянии красноярского воздуха, их визуализацию и рассказ о факторах изменения экологической ситуации в Красноярске с привлечением мнения экспертов.

Актуальность проекта заключается в том, что данных об изучаемой нами проблеме достаточно много, и на региональном уровне в формате журналистики данных вопрос красноярской экологической ситуации еще не изучался.

Данные в проекте представлены в различных форматах: статичные и интерактивные инфографики, фотографии, видео, интервью.

Аудитория проекта – жители города Красноярска.

2.1 Подготовительный этап работы над проектом

На данном этапе мы определили актуальную для красноярской аудитории тему – загрязнение воздуха. Мы составили план работы над проектом и определили три ключевых аспекта, которые в последующем нужно было подкрепить данными:

- 1 Основные факторы ухудшения состояния воздуха;
- 2 Основные источники загрязнения;
- 3 Возможные варианты решения проблемы

На этом же этапе мы определили название проекта – «Задержите дыхание».

2.2 Сбор данных

На этом этапе мы определили источники данных.

За время работы мы обратились к базам данных Правительства Красноярского края, министерства экологии края, Красноярскстата, Среднесибирского УГМС, базам данных независимого проекта «Красноярск. Небо». Кроме того, нашим источником информации были эксперты, объясняющие на примере данных закономерности определенных явлений. Все данные обрабатывались в таблицах Excel.

Таблица 1 – Источники данных для проекта «Задержите дыхание»

База данных	Источник данных
Данные о количестве дней в режиме НМУ в Красноярске за период с 2010 по 2018 год.	Сайт администрации Красноярска
Данные об объеме выбросов от	Красноярскстат

стационарных источников в Красноярске.	
Данные о доле источников загрязнений в Красноярске	Сайт правительства края
Данные об источниках выбросов вредных веществ в атмосферный воздух Красноярска	Сайт правительства края, Государственный доклад о состоянии и охране окружающей среды в Красноярском крае за 2016 год
База данных	Источник данных
Таблицы загрязнений по разным веществам с датчиков за период с октября 2017-го по апрель 2018-го	Сайт министерства экологии
Данные о превышениях ПДК с датчиков, установленных сайтом Среднесибирского УГМС	Отчеты СУГМС с сайта ведомства, а также Государственный доклад о состоянии и охране окружающей среды в Красноярском крае за 2016 год
Базы данных с датчиков активистов за 2017 год	Проект «Красноярск.Небо»

Обратившись к открытым данным краевого правительства о количестве дней в режиме НМУ за период с 2010 по 2018 год, мы обнаружили, что они содержатся в нетабличном текстовом формате, и информация указана за каждый месяц определенного года. Мы вручную привели их в табличный вид: зафиксировали, сколько за каждый месяц было объявлено таких дней, после этого и суммировали данные, получив результат за год.

Также мы запросили базы данных активистов проекта «Красноярск. Небо», и получили 11 таблиц с данными с датчиков из разных районов города. Однако мы столкнулись с тем, что замеры выбросов частиц PM_{2,5} по каждому датчику начинались не с начала года, а за разные месяцы. Мы отобрали три базы данных с наиболее схожим периодом замера и взяли показатели за полгода – с октября 2017-го года по апрель 2018-го. Принципом отбора также стало место нахождения датчика: мы взяли данные с ул. Ады Лебедевой, ул.

Копылова, ул. Юности – условно говоря, с центра, Октябрьского района и правого берега соответственно.

За каждый день были представлены почасовые наблюдения, и нам потребовалось определить средний показатель за день – поэтому мы вручную в таблицах Excel привели к среднему числу цифры за каждый день. После этого мы создали единую базу данных по всем трём датчикам.

	A	B	C	D	E	F	G	H	I	J	K
1	date	pm25									
2	17.10.2017	15									
3	18.10.2017	12									
4	19.10.2017	20									
5	20.10.2017	19									
6	21.10.2017	10									
7	22.10.2017	23									
8	23.10.2017	53									
9	24.10.2017	12									
10	25.10.2017	0									
11	26.10.2017	0									
12	27.10.2017	0									
13	28.10.2017	2									
14	29.10.2017	6									
15	30.10.2017	32									
16	31.10.2017	17									
17	01.11.2017	154									
18	02.11.2017	240									
19	03.11.2017	12									
20	04.11.2017	7									
21	05.11.2017	5									
22	06.11.2017	8									
23	07.11.2017	11									
24	08.11.2017	5									
25	09.11.2017	3									

Рисунок 1 – База данных проекта «Красноярск.Небо»

Также для создания карты выбросов, зафиксированных датчиками Минэкологии края и Среднесибирским УГМС, мы взяли данные с сайтов этих ведомств и из Государственного доклада о состоянии и охране окружающей среды в Красноярском крае за 2016 год.

Сначала мы создали базу данных, основанную на таблицах министерства экологии. Всего ведомство контролирует выбросы на шести точках, одну из них мы исключили: датчик «Кубеково» ввиду его дальности от города. Далее мы внесли в таблицы Excel данные по каждому датчику и сделали колонки с разными веществами, после чего внесли показатели за период с октября прошлого года по апрель 2018-го и вывели среднее число превышения предельно допустимой концентрации по каждому веществу за полгода. После нанесения на карту стало ясно, что этих данных мало: большинство датчиков находится практически в крайних точках города, и не охватывают, например, центр Красноярска, Октябрьский район. При этом Среднесибирское УГМС имеет больше датчиков, но делает замеры другим способом и делится данными

в нетабличном, обзорном виде, что делает выборку разрозненной. Но, чтобы наша карта загрязнений была более полной, мы собрали данные с датчиков СУГМС из отчетов, выложенных на сайте, а также из Государственного доклада о состоянии и охране окружающей среды в Красноярском крае за 2016 год, как мы упомянули выше. Работа с картой вызвала большие сложности ввиду того, что данные разрозненные, замеры делаются по-разному. К сожалению, это мы определили на этапе, когда решение сделать карту уже было принято и для карты мы определили логическое место в тексте. Такой порядок работы с данными был ошибкой, поскольку практики журналистики данных просят делать с точностью обратное: сначала определить данные, собрать их, а уже после основывать на них тезисы/вставлять в проект.

A	B	C	D	E	F	G	H	I	J	K	L	M
Северный	оксид угл	оксид азот	диоксид азот	диоксид серы	гидрофторид	сероводород	фенол	гидрохлорид	диоксид серы	формальдегид		
мар.18	1,04	1,2	1,26	1,21	1,90	1,75	1,3	1,6	0,25			
фев.18	1,30	1,51	0,77	1,61	4,15	1,00	2,0	1,6	0,29			
янв.18	1,26	1,68	0,69	1,76	1,30	2,13	0,4	1,6	0,33			
дек.17	2	1	1	1		3	1	10	1			
ноя.17	1	2	3			1						
окт.17	1	2		2	1	1,776	1		3			
Среднее:	1,3	1,6	1,3	1,5	2,1	2	1	3,7	1			
Солнечный												
мар.18	0,8	0,25	0,57	0,94	4,85	3,25	0,1	1,8	1,11			
фев.18	2,18	0,45	0,61	1,97	5,95	8,38	0,2	1,65	3,02			
янв.18	1,42	0,45	0,71	1,26	1,75	8,13	0	1,65	1,15			
дек.17	2	1	1	1					3			
ноя.17	1	2	1									
окт.17	1,5	1	1	1	1	2		3				
Среднее	2	1	0,8	1	3,4	5,4	0,2	2	2			
Черемушки												
мар.18	0,78	0,88	3,15	0,91	3,85	1	0	1,75	0,13	0		
фев.18	0,92	1,34	0,74	1,72	6,7	1,5		1,75	0,1	5,8		
янв.18	1,34	1,31	2,88	1,32	0	0		0	0,23	3,86		
дек.17	1	1	2	1	0	2		1	0,2	0		
ноя.17	0	0	0	0	0	1,125		0		0		
окт.17	1	2	0	0	0	0	0	0		0		
Среднее	0,8	1,1	1,5	0,8	5	2		2	0,2	1,9		

Рисунок 2 – База данных с сайта министерства экологии

С сайта Красноярскстата мы взяли информацию об объеме выбросов от стационарных источников загрязнений с 2010 года по настоящее время, а также об источниках загрязнений по ряду веществ. Эти открытые данные использовали как фактическую опору для текста проекта.

Источниками информации для нас стали также эксперты и обычная красноярка. Первое интервью мы сделали с директором Института экологии и географии СФУ Русланом Шарафутдиновым. С одной стороны, он выступил экспертом, разъясняющим закономерность, которую мы нашли на основе больших данных, а именно: по данным Красноярскстата, объем выбросов

снижается, при этом количество дней НМУ растет, по данным с сайта правительства края. Кроме того, директор института сам стал источником, объяснив и другие факторы ухудшения состояния экологии.

Второе интервью – с краснояркой Ольгой – изначально не планировалось. После того, как появилась канва проекта, стало ясно, что на фоне данных не хватает человеческой истории. Таким образом, история героини стала вводной частью, приближающей проблему к читателю.

Еще одним экспертом стал Александр Сосновский - инспектор отдела экологического надзора министерства экологии края. Интервью с ним – взгляд изнутри на то, как действует в сложившейся ситуации министерство.

2.3 Этап визуализации

Проект «Задержите дыхание» представлен на платформе Tilda. Этот сервис мы выбрали, потому что он предлагает много решений для визуализации и представления информации. Проект сделан в формате лонгрида - непрерывного текста с мультимедийными вставками. На этом этапе использовались такие мультимедийные элементы, как инфографика (интерактивные линейные графики и статичные диаграммы), интерактивная карта, фотогалереи, два видеофрагмента и аудиодорожка.

Фотография на обложке проекта, сделанная автором лонгрида, имеет метафорическое значение и перекликается по смыслу с названием: на фото изображена девушка в противогазе, пришедшая на митинг.

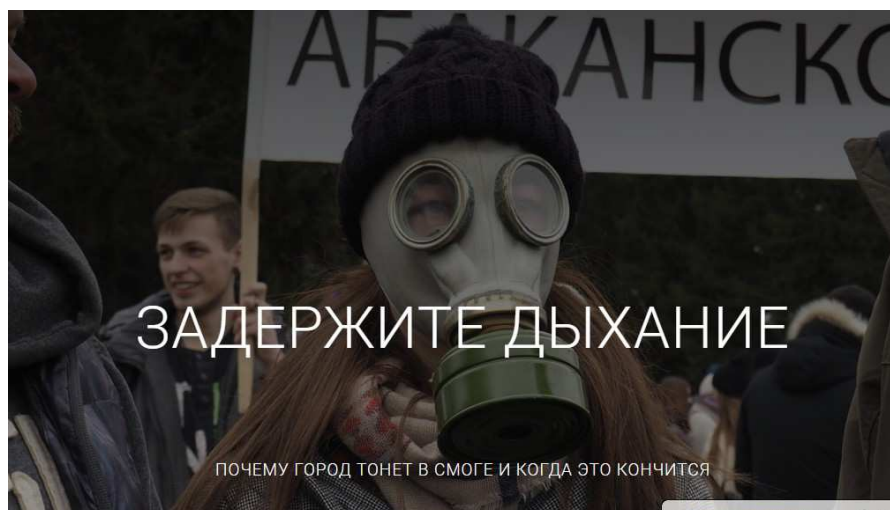


Рисунок 3 – Обложка проекта «Задержите дыхание»

Текст начинается с истории красноярки, у которой обостряется аллергия на выбросы в дни НМУ. Главным образом, в этой части повествования присутствует интервью о её недуге, иллюстрируемое фотогалереей и 1-минутным видеороликом, сделанными в ходе разговора с ней.

В следующей части проекта под названием «Ну почему опять НМУ?» мы представляем читателю закономерность, обнаруженную нами в данных Красноярскстата и правительства края. Для представления данных о количестве дней НМУ за 8 лет мы создали интерактивный линейный график, показывающий рост количества таких дней. Для визуализации мы использовали сервис Piktochart, поскольку он не только удобен в использовании, но и предоставляет возможность наложить фотографию (в нашем случае – пыльного Красноярска) на фон графика.



Рисунок 4 – Интерактивный линейный график «Дни НМУ в Красноярске»

В этой же части мы вставили элементы интервью эксперта с объяснением обнаруженных нами закономерностей, и часть его слов представлена видеороликом длительностью 1:23 минуты.

В следующей части, рассказывающей об источниках загрязнений, мы также использовали данные Красноярскстата. Изначально визуализация была сделана с помощью Piktochart в том же стиле, что и линейный график; мы создали пять круговых статичных диаграмм. Однако позднее оказалось, что данные об источниках загрязняющих веществ лучше представить не диаграммами, а столбчатыми графиками, которых в Piktochart нет. При создании диаграммы и графиков мы решили использовать ресурс Infogram, который позволяет создавать этот вид графиков.

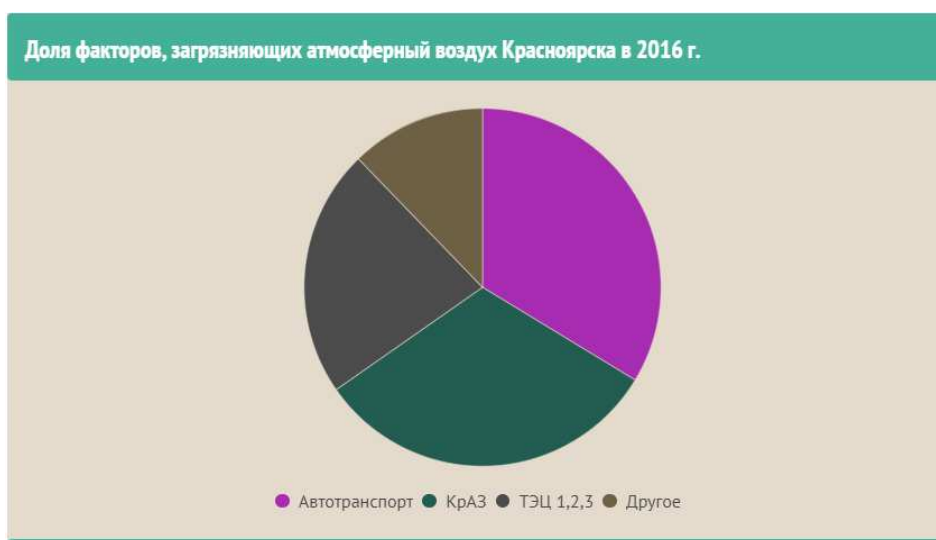


Рисунок 5 – Круговая интерактивная диаграмма «Источники загрязнения воздуха в Красноярске»

Таким образом, мы пришли к выводу, что данные об источниках загрязнений с сайта министерства экологии лучше будет для сравнения по веществам представить в виде интерактивных столбчатых диаграмм.



Рисунок 6 – Интерактивный столбчатый график «Источники загрязняющих веществ в Красноярске»

Для визуализации данных с 13 датчиков министерства экологии мы использовали встроенные в Tilda «Яндекс карты». Нужно сказать, что данный интерфейс не позволяет добавлять свои метки и цвета – нужно использовать только предложенные сервисом. Но поскольку встроенные карты «Яндекса» достаточно просты и нет сложности с их встраиванием – мы воспользовались ими.



Рисунок 7 – Карта загрязнений с датчиков минэкологии

Кнопкой, ведущей на отдельную страницу, мы перенаправили читателя на цифры с датчиков активистов проекта «Красноярск. Небо». Там можно увидеть интерактивный линейный график загрязнений по дням за полгода. График создан с помощью сервиса Tableau Public. Выбор обусловлен тем, что

это приложение во многом интуитивно понятно пользователю – в работе с большими данными это полезное свойство. Кроме того, оно имеет «умный» подход к обработке больших данных и предлагает разные способы визуализации в зависимости от предложенных данных. Уже после завершения этапа визуализации, во время встраивания графика на сайт, мы столкнулись с тем, что сервис не позволяет регулировать высоту графика, предлагая только сделать его длиннее/короче и выше/ниже. Длинный, растянутый график удобнее для чтения, однако сервис оставляет пустое место после графика. В связи с этим мы поместили эти данные на отдельную страницу, и дали возможность пользователю, нажав кнопку, вернуться обратно на главную страницу.

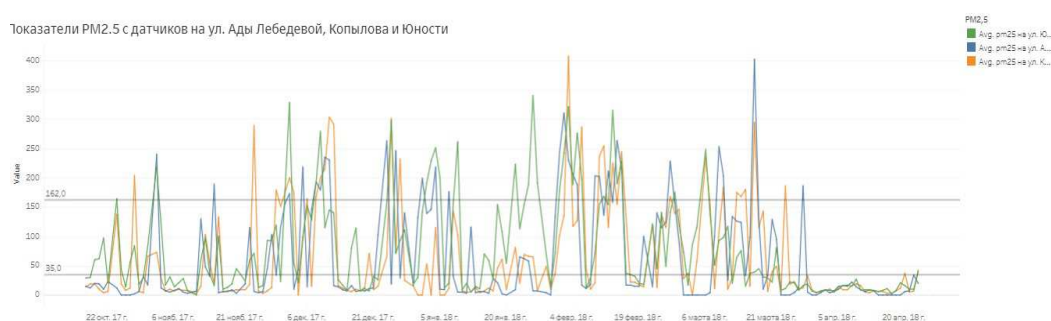


Рисунок 9 – График загрязнений с датчиков «Красноярска.Небо»

В последней, заключительной части, из мультимедийных элементов использованы фотогалерея и аудио. Галерея – репортаж с митинга, и она эмоционально подкрепляет текст о том, что делать Красноярску с загрязнением воздуха. Ту же функцию выполняет и аудиовставка с песней о красноярском воздухе.

Последний элемент лонгрида – «подвал». В нём мы указали авторство, год создания проекта и привели ссылку на Институт филологии и языковой коммуникации. Также в этом блоке указаны источники, использованные в проекте.



© 2018 Задержите дыхание. Текст, фото, видео: Ольга Ясева.
Монтаж видео: Любовь Полежаева.
Источники: министерство экологии Красноярского края, официальный портал Красноярского края, Среднесибирское УГМС, базы данных проекта "Красноярск.Небо".

Рисунок 10 – «Подвал» проекта

2.4 Этап публикации

Проект был опубликован в Независимом информационном агентстве от 14.06.2018. Режим доступа:

На этом этапе не возникло сложностей, поскольку на момент создания проекта автор работал в данном СМИ.

2. 5 Анализ работы над проектом «Задержите дыхание»

Идея создать проект о состоянии красноярского воздуха исходит из актуальности этой проблемы для жителей города – потенциальной аудитории проекта. В городском медиаполе эта тема часто освещается, однако чаще всего в русле разговора с экспертами и обращения к общим данным. Мы решили посмотреть на важную для красноярцев тему немного под другим углом – с использованием больших данных. Это уникальный опыт, поскольку баз данных для исследования много, но из коллег в городе этим мало кто занимается в виду времязатратности. Изначально мы предполагали, что полноценного проекта журналистики данных не получится, поскольку сбором информации, работой с базами данных и визуализацией занимается один человек и в небольшие сроки. Кроме того, это первый опыт работы с большими данными у автора.

Первым сложным моментом стало определение структуры проекта. Было сложно определить, каким образом осветить такую масштабную тему. Поскольку проект журналистики данных несколько отличается от обычного журналистского проекта упором на данные, важно было выбрать правильную стратегию. Мы решили сделать пошаговый разбор вопроса об экологической ситуации. Сначала описали проблему, а затем разбили её на части: факторы ухудшения состояния воздуха, основные источники загрязнений и способы решить проблему. Далее мы либо приводили данные и позже объясняли на их примере закономерности, либо приводили тезис и подтверждали его данными.

Структура проекта менялась не один раз. Во-первых, как мы упоминали в параграфе об источниках данных, изначально часть с историей красноярки Ольги об аллергии не предполагалась, но она органично вписалась в проект и после написания двух других частей. Во-вторых, идея создать карту загрязнений на основе официальных датчиков сначала была отмечена из-за разнородности данных – в итоге карта все-таки появилась.

В пособии «Думай, как журналист» подчёркивается, что у структуры проекта журналистики данных есть своя специфика. Мы выбрали так называемую структуру «Кебаб», в которой история начинается с человека, а продолжается данными. Такой ход мы считаем удачным, поскольку проблематика проекта в первую очередь касается каждого жителя города в частности.

На этапе сбора информации мы столкнулись с тем, что некоторые данные, обнародованные правительством края, выложены в нетабличном виде, из-за чего пришлось создавать отдельные таблицы с такими данными. Кроме того, при создании карты загрязнений, мы обнаружили, что не все данные выложены в открытом доступе. Так, у СУГМС и министерства экологии по-разному обрабатывается информация с датчиков и по-разному представляются отчеты о замерах. Трудность была и в том, чтобы разобраться в таблицах

В этом случае можно сделать вывод, что работа с базами данных в министерстве отлажена недостаточно хорошо, по крайней мере, в отношении

доступности этих данных журналистам. В то же время независимые «контролёры» воздуха поделились базами данных в короткий срок, и базы данных были удобном для работы виде.

Отсутствие опыта работы с большими данными в таблицах Excel стало одной из главных препятствий, которые в работе над проектом предстояло преодолеть. Так, мы научились приводить большое количество данных к среднему числу (несмотря на то, что это простая операция, изначально мы ее не догадались использовать, а она оказалась полезной). Кроме того, предстояло разобраться, как эти таблицы вставлять в приложения для визуализации: возникали проблемы с нечитаемостью форматов, например.

Трудно было визуализировать данные активистов «Красноярск.Небо». Было непросто, поскольку данные по каждому датчику начинаются с разных дат, и для репрезентативности выборки. Большинство датчиков фиксировали информацию с ноября 2017 года, поэтому мы взяли этот месяц. Мы выбрали три базы данных с трех датчиков в разных районах города, и по нашему мнению, этой информации мало. При этом если брать больше информации, график получается чрезмерно массивным и непонятным. Поэтому мы не стали увеличивать число баз данных. Еще одной трудностью стала невозможность встроить сделанный в Tableau Public график так, чтобы это не портило эстетику проекта: то график «уплывал» вправо, то появлялось белое пространство снизу, которое нельзя было убрать. Поэтому для того, чтобы график было удобнее рассматривать и он не портил пространство сайта, мы перенесли интерактивную его версию на другую страницу, а на главной оставили скриншот. Это тоже создает некоторые неудобства для пользования, потому что в Tilda у кнопки нет возможности перенаправления на блок другой страницы. В целом это приложение оказалось очень полезным в работе с большими данными: оно предлагает способы визуализации предложенных тобой таблиц – и это положительная сторона приложения. С подобным интерфейсом мы столкнулись впервые, кроме того, в приложении все действия указаны на английском языке, и это несколько усложнило работу вначале. Но

после нескольких попыток создать график, стало интуитивно понятно, как использовать Tableau Public. К сожалению, у этих таблиц возникли проблемы при встраивании в Tilda, и о них мы писали ранее. Мы связываем тот факт, что пришлось перенести интерактивный график, сделанный в Tableau, на другую страницу с неопытностью в работе с подобными приложениями и в целом с первым опытом программирования и верстки. Навык программирования и работы с таблицами полезен, когда нужно начать работу с данными – без него работа над проектом сильно тормозится и усложняется, так произошло и в нашем случае. Большая часть времени работы над лонгридом ушла на то, чтобы разобраться, как обработать таблицы и визуализировать их, чем на создание контента и верстку.

В работе мы использовали два подхода к данным. В одном случае большие данные становились инфоповодом (как в случае с сопоставлением дней НМУ и объема выбросов от стационарных источников за несколько лет) и для их объяснения мы обращались к эксперту – директору Института экологии и географии СФУ. Цель такого объясняющего интервью, по нашему мнению, достигнута.

В другом случае большими данными (графиком с цифрами проекта «Красноярск. Небо») мы иллюстрировали тезис «данные о загрязнении воздуха, предоставленные министерством экологии и общественников сильно расходятся». В этом случае мы использовали так называемое интервью «для привлечения к ответственности» – то есть разговор с официальным лицом о том, почему данные расходятся и о ситуации с воздухом в целом.

В целом, сказать, что наш проект полностью выступает в области журналистики данных, нельзя. Скорее, это мультимедийный проект с применением журналистики данных и вставками данных, потому что на первый план вышли мнения о проблеме загрязнения, а не оценка найденных данных. Мы это связываем с отсутствием опыта работы с большими данными и таблицами. Вероятно, стоило сначала собрать все имеющиеся данные и на их

основе вывести гипотезы, которые по ходу текста нужно было бы подтверждать или опровергать.

В работе нам не хватало навыков программирования и опыта в работе с большими данными – это оказалось сложно и времязатратно. Поскольку над проектом трудился один человек, на создание, например, дополнительных инфографик или текста не хватило времени. Но этот опыт показал, что, даже впервые столкнувшись с большими данными, можно постепенно понять, как их применить.

Несмотря на вышеперечисленные недостатки, мы считаем, что проект получился, поскольку он достиг цели – ответить на вопросы о красноярском воздухе с использованием больших данных, как это и задумывалось изначально. Кроме того, наш опыт доказал, что можно создавать лонгрид с применением больших данных и в региональной журналистике.

Подводя итог, скажем: лонгрид «Задержите дыхание» – мультимедийный проект с использованием больших данных, созданный на базе Tilda. Режим доступа: <http://zaderzhite.dyhanie.tilda.ws/1>

В проекте представлены:

1 Интерактивный линейный график, созданный в Piktochart. Режим доступа: <https://create.piktochart.com/output/29728117-new-piktochart>;

2 Пять статичных круговых диаграмм, также выполненных в Piktochart и преобразованных в фотогалерею;

3 Интерактивная Яндекс карта, встроенная в Tilda;

4 Интерактивный линейный график, сделанный в Tableau. Режим доступа:

https://public.tableau.com/views/fitwidth22/Sheet2?:embed=y&:display_count=yes;

5 Три фотогалереи;

6 Две видеозаписи;

7 Одна аудиозапись

Тема проекта актуальна для жителей города, при этом он не привязан к определенному информационному поводу, соответственно, его можно читать спустя время после публикации.

ЗАКЛЮЧЕНИЕ

Журналистика данных – это новое направление, которое активно развивается пока только за рубежом. Так, проекты на основе больших данных можно встретить в Guardian и New York Times. В российской практике журналистика данных также встречается, но реже и чаще на уровне федеральных СМИ. В то время как для региональной журналистики работа с большими данными – пока редкость.

Цель проекта состояла в том, чтобы создать авторский мультимедийный проект на основе больших данных. В итоге мы разработали проект о состоянии воздуха в Красноярске «Задержите дыхание». Прделанная работа соответствует определенным нами вначале задачам и цели.

В теоретической части работы мы дали определения понятиям «большие данные», «открытые данные» и «журналистика данных», а также описали этапы работы над проектом журналистики данных.

Мы выяснили, что журналистика данных – это работа с большими базами данных, при которой от журналиста требуются навыки в том числе программиста, социолога и дизайнера. Работа над проектом с использованием больших данных включает в себя несколько основных этапов: подготовка, сбор информации, обработка собранных данных, этап визуализации и публикации.

Также мы выяснили, что проект журналистики данных схож с обычным лонгридом, так как он тоже создается в режиме онлайн, включает в себя мультимедийные и интерактивные элементы.

Проект «Задержите дыхание», созданный в рамках исследования, - это лонгрид на площадке Tilda, рассказывающий с использованием больших данных о том, почему в Красноярске ухудшается состояние воздуха, как с этим борются в городе правительство и активисты и какое есть решение у этой проблемы.

Во время работы над проектом мы использовали открытые данные правительства Красноярского края, министерства экологии Красноярского края, администрации Красноярска, Среднесибирского УГМС, Красноярскстата. Также мы запросили данные проекта «Красноярск.Небо» и обратились к экспертам с вопросами о состоянии красноярского воздуха. На основе этих данных мы создали свои базы данных и визуализировали их.

На этапе визуализации мы создали два интерактивных линейных графика, одну интерактивную круговую диаграмму, один интерактивный столбчатый график, одну интерактивную карту, два видеоролика, две фотогалереи и одну аудиозапись.

По нашему мнению, наиболее удачным проект получился с точки зрения журналистики: мы удачно выбрали героев и рассказали о красноярском воздухе с разных точек зрения. При этом в полной мере назвать проект «Задержите дыхание» журналистикой данных нельзя, поскольку с большими данными мы работали впервые и не смогли в полной мере рассчитать силы, поэтому

визуализировать данные с проекта «Красноярска.Небо» в том виде, как хотелось бы, не вышло.

Создав проект с использованием больших данных, мы доказали, что региональные журналисты так же, как федеральные и зарубежные специалисты, способны создавать масштабные лонгриды на основе баз данных. Подобные проекты позволяют глубже погрузиться в социально важные вопросы и исследовать проблемы, важные для жителей.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Бегтин, И. В. Готовы ли мы к журналистике данных? [Электронный ресурс] / И. В. Бегтин // Полит.ру. – 2013. – Режим доступа: http://polit.ru/article/2013/04/29/data_journalism.

2 Бегтин, И. В. В России нет системного подхода к журналистике данных [Электронный ресурс] / И. В. Бегтин // РИА Новости. – 2013. – Режим доступа: <https://ria.ru/interview/20131107/975231213.html>.

3 Булаева, М. Н. Мультимедийный лонгрид как новый журналистский формат [Электронный ресурс] / М. Н. Булаева // Журналистский ежегодник. – 2015. – №4. – Режим доступа: <https://cyberleninka.ru/article/n/multimediynyy-longrid-kak-novyy-zhurnalistskiy-format>.

- 4 Голомбински, К., Добавь воздуха! Основы визуального дизайна для графики, веб и мультимедиа / К. Голомбински, Р. Хаген. – СПб: Питер, 2013. – 130 с.
- 5 Доброва, И. Инфографика: шаг вперед, два назад? [Электронный ресурс] / И. Доброва // InfoGrapher. – 2016. – Режим доступа: <http://infographer.ru/infografika-shag-vpered-dva-nazad/>.
- 6 Доброва, И. Три уровня визуализации: данных, информации, знания [Электронный ресурс] / И. Доброва // InfoGrapher. – 2015. – Режим доступа: <http://infographer.ru/3dimentions/>.
- 7 Константарас, Э. Думай, как дата-журналист [Электронный ресурс] / Э. Константарас // – 2013. – Режим доступа: <http://www.odecanet.org/data-journalism-manual>.
- 8 Жолудь, Р. В. Журналистика данных: предпосылки возникновения, функции, возможности [Электронный ресурс] / Р. В. Жолудь // Вестник воронежского государственного университета. – № 4. – 2013. – С.23-27. – Режим доступа: <https://elibrary.ru/item.asp?id=22871877>.
- 9 Клеменков, П. А., Большие данные: современные подходы к хранению и обработке [Электронный ресурс] / П. А. Клеменков, С. Д. Кузнецов // Труды ИСП РАН. – 2012. – №.2 – С.143-155. – Режим доступа: <https://cyberleninka.ru/article/n/bolshie-dannye-sovremennyye-podhody-k-hraneniyu-i-obrabotke>.
- 10 Лазутина, Г. В. Основы творческой деятельности журналиста / Г. В. Лазутина // М.: МедиМир, 2008. – 240 с.
- 11 Левченко, В. Ю., Big Data, Open Data, Linked Data, метаданные в PR: актуальные модели трансформации теории и практики [Электронный ресурс] / В. Ю. Левченко, М. Г. Шилина // Медиаскоп. – 2003-2013. – Режим доступа: <http://www.mediascope.ru/node/1486#14>.
- 12 Лукина, М. М. Интернет – СМИ: Теория и практика: Учеб. пособие для студентов вузов / М. М. Лукина. – М.: Аспект Пресс, 2010. – 348 с.

- 13 Лукина, М. М., СМИ в пространстве Интернета / М. М. Лукина, И. Д. Фомичева // Сер. «Интернет-журналистика». – Вып. 1. М., 2005. – 55 с.
- 14 Машкова, С. Г. Интернет-журналистика: учебное пособие / С. Г. Машкова // – Тамбов, 2006. – 79 с.
- 15 Майер-Шенбергер, В. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / В. Майер-Шенбергер, К. Кукьер. – Москва: Манн, Иванов и Фербер, 2014. – 240 с.
- 16 Моррисон, А. Поиск эффективных инструментов работы с большими данными / А. Моррисон // Технологический прогноз. – 2010. – № 3. – С. 41-50.
- 17 Найдич, А. Big Data: проблема, технология, рынок [Электронный ресурс] / А. Найдич // Компьютер Пресс. – 1999-2014. – С. 25 – Режим доступа: <http://compress.ru/article.aspx?id=22725>.
- 18 Нефедьева, К. В. Инфографика – визуализация данных в аналитической деятельности [Электронный ресурс] / К. В. Нефедьева // Труды СПбГИК. – 2013. – С. 89-93. – Режим доступа: <https://cyberleninka.ru/article/n/infografika-vizualizatsiya-dannyh-v-analiticheskoy-deyatelnosti>.
- 19 Панюкова, С. А. Роль открытых данных в развитии журналистики данных [Электронный ресурс] / С. А. Панюкова // Знак: проблемное поле медиаобразования. – 2015. – №1 (15). – Режим доступа: <https://cyberleninka.ru/article/n/rol-otkrytyh-dannyh-v-razviii-zhurnalistiki-dannyh>.
- 20 Пособие по журналистике данных [Электронный ресурс] // Под ред. Д. Грей. М.: РИА Новости, 2013. – 240 с. – Режим доступа: https://ria.ru/files/book/_site/.
- 21 Романенко, Е. В. Место Big Data в современной социально-экономической жизни общества [Электронный ресурс] / Е. В. Романенко // Инновационная наука. – 2016. – № 4-3 (16). – С.143-144. – Режим доступа: <http://cyberleninka.ru/article/n/mesto-big-data-v-sovremennoy-sotsialno-ekonomicheskoy-zhizni-obschestva>.

22 Симакова, С. И. Журналистика данных как современное направление журналистики / С. И. Симакова // Научный журнал «Знак»: проблемное поле медиаобразования. – 2013. – №1 (11). – С. 52-56.

23 Симакова, С. И. Журналистика данных как современное направление журналистики [Электронный ресурс] / С. И. Симакова // Знак: проблемное поле медиаобразования. – 2013. – №1 (11). – С. 34-45. – Режим доступа: <https://cyberleninka.ru/article/n/zhurnalistika-dannyh-kak-sovremennoe-napravlenie-zhurnalistiki>

24 Симакова, С. И. Инфографика как средство визуализации экономической информации в СМИ [Электронный ресурс] / С. И. Симакова // Вестник ЧелГУ. – 2014. – №23 (352). – Режим доступа: <https://cyberleninka.ru/article/n/infografika-kak-sredstvo-vizualizatsii-ekonomicheskoy-informatsii-v-smi>.

25 Шилина, М. Г., Big Data, Open Data, Linked Data, метаданные в PR: актуальные модели трансформации теории и практики [Электронный ресурс] / М. Г. Шилина, В. Ю. Левченко // Журнал Факультета журналистики МГУ имени М. В. Ломоносова «Медиаскоп». – 2014. – №1. – Режим доступа: <http://www.mediascope.ru/node/1486>.

26 Шилина, М. Г. Data Journalism – дата-журналистика, журналистика метаданных – в структуре медиакоммуникации: к вопросу формирования теоретических исследовательских подходов [Электронный ресурс] / М. Г. Шилина // Журнал Факультета журналистики МГУ имени М. В. Ломоносова «Медиаскоп». – 2013. – №1. – Режим доступа: <http://www.mediascope.ru/node/1263>.

27 Шилина, М. Г. Медиавизуалистика в парадигме big data и virtual reality / М. Г. Шилина // Медиаальманах. – 2017. – № 1. – С. 12–24.

28 Шилина, М. Г. Медиакоммуникация: тенденции трансформации. Новые парадигмы исследований массовых коммуникаций [Электронный ресурс] / М. Г. Шилина // Медиаскоп. – 2009. – № 3. – Режим доступа: <http://www.mediascope.ru/node/404>.

29 Шилина, М. Г. Медиа XXI века как объект исследования (Несколько тезисов об инновационных подходах к изучению массовых коммуникаций) [Электронный ресурс] / М. Г. Шилина // Медиаальманах. – 2009. – №. 5. – С. 6-12. – Режим доступа: http://www.mediaalmanah.ru/files/59/2013_6_7_shilina.pdf.

30 Шерстюкова, М. Н. Дата-журналистика как новое направление в системе средств массовой коммуникации [Электронный ресурс] / М. Н. Шерстюкова // Международный электронный научно-образовательный журнал «Медиа. Информация. Коммуникация». – 2012. – №1. – Режим доступа: <http://mic.org.ru/1-nomer-2012/65-1-shestukova-2>

31 Goranson, C. Data Visualization For Big Data [Электронный ресурс] / С. Goranson // 2014, pp.7-14.

32 The Guardian Datablog [Электронный ресурс] // – 2009. – Режим доступа: <http://www.theguardian.com/news/datablog>.

19

Федеральное государственное автономное
образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Институт филологии и языковой коммуникации
Кафедра журналистики

УТВЕРЖДАЮ
Заведующий кафедрой
К. В. Анисимов
« 18 » июня 2018г.

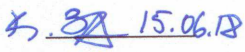
БАКАЛАВРСКАЯ РАБОТА

42.03.02 Журналистика

**ОПЫТ ПРИМЕНЕНИЯ ЖУРНАЛИСТИКИ ДАННЫХ В АВТОРСКОМ
МУЛЬТИМЕДИЙНОМ ПРОЕКТЕ**

Руководитель  15.06.18 ст.преподаватель Д.А. Устюжанина

Выпускник  15.06.18 О.А. Ясева

Консультант  15.06.18 канд.филос.наук, К.А. Зорин
доцент

Нормоконтролер  15.06.18 ст.преподаватель О.В. Богуславская

Красноярск 2018